

WHAT IS CLAIMED IS:

1. A text sentence comparison method comprising:
converting a first text sentence and a second text sentence
5 into a first R tree and a second R tree, respectively;
calculating a distance between the first R tree and the
second R tree on the basis of a distance between two R trees,
which is defined at least in accordance with a condition of
a mapping between vertexes and edges of the two R trees; and
10 calculating a distance between the first text sentence
and the second text sentence on the basis of the calculated
distance between the first R tree and the second R tree, wherein:
in the conversion,
word information contained in the first text
15 sentence are allotted to the vertexes of the first R tree;
word information contained in the second text
sentence are allotted to the vertexes of the second R
tree;
case information contained in the first text
20 sentence are allotted to the edges of the first R tree;
and
case information contained in the second text
sentence are allotted to the edges of the second R tree.
- 25 2. The text sentence comparison method according to

claim 1, wherein in the calculation of the distance between the first R tree and the second R tree:

a mapping weight of a mapping from the first R tree to the second R tree;

5 a distance between a forest, which the first R tree includes, and a forest, which the second R tree include; and

a distance between a subtree, which the first R tree includes, and a subtree, which the second R tree includes; are calculated.

10

3. The text sentence comparison method according to claim 2, wherein:

the mapping weight is calculated on the basis of a word substitution weight, a word deletion weight, a word insertion
15 weight, a case substitution weight, a case deletion weight, and a case insertion weight.

4. The text sentence comparison method according to claim 1, further comprising:

20 setting the condition of the mapping between the first and second R trees.

5. The text sentence comparison method according to claim 1, wherein the condition of the mapping between the first
25 and second R trees includes:

the mapping is a one-to-one mapping;
the mapping preserves parent-child relationship;
the mapping preserves structure; and
the mapping between the vertexes does not intersect with
5 the mapping between the edges.

6. The text sentence comparison method according to
claim 1, further comprising:

inputting the first text sentence and the second text
10 sentence; and

outputting the calculated distance between the first text
sentence and the second text sentence.

7. A text sentence comparison method comprising:
15 converting a first text sentence and a second text sentence
into a first RO tree and a second RO tree, respectively;

calculating a distance between the first RO tree and the
second RO tree on the basis of a distance between two RO trees,
which is defined at least in accordance with a condition of
20 a mapping between vertexes and edges of the two RO trees; and

calculating a distance between the first text sentence
and the second text sentence on the basis of the calculated
distance between the first RO tree and the second RO tree,
wherein:

25 in the conversion,

word information contained in the first text sentence are allotted to the vertexes of the first RO tree;

5 word information contained in the second text sentence are allotted to the vertexes of the second RO tree;

case information contained in the first text sentence are allotted to the edges of the first RO tree; and

10 case information contained in the second text sentence are allotted to the edges of the second RO tree.

8. The text sentence comparison method according to claim 7, wherein in the calculation of the distance between
15 the first RO tree and the second RO tree:

a mapping weight of a mapping from the first RO tree to the second RO tree;

a distance between a forest, which the first RO tree includes, and a forest, which the second RO tree include; and

20 a distance between a subtree, which the first RO tree includes, and a subtree, which the second RO tree includes; are calculated.

9. The text sentence comparison method according to
25 claim 8, wherein:

the mapping weight is calculated on the basis of a word substitution weight, a word deletion weight, a word insertion weight, a case substitution weight, a case deletion weight, and a case insertion weight.

5

10. The text sentence comparison method according to claim 7, further comprising:

setting the condition of the mapping between the first and second RO trees.

10

11. The text sentence comparison method according to claim 7, wherein the condition of the mapping between the first and second RO trees includes:

the mapping is a one-to-one mapping;

15

the mapping preserves parent-child relationship;

the mapping preserves brother relationship;

the mapping preserves structure; and

the mapping between the vertexes does not intersect with the mapping between the edges.

20

12. The text sentence comparison method according to claim 7, further comprising:

inputting the first text sentence and the second text sentence; and

25

outputting the calculated distance between the first text

sentence and the second text sentence.

13. A text sentence comparison apparatus comprising:
an input section for inputting a first text sentence and

5 a second text sentence;

a tree structure conversion section for converting the
first text sentence and the second text sentence into a first
R tree and a second R tree, respectively;

a distance calculation section for calculating a distance
10 between the first R tree and the second R tree on the basis
of a distance between two R trees, which is defined at least
in accordance with a condition of a mapping between vertexes
and edges of the two R trees; and

a semantic content comparison section for calculating
15 a distance between the first text sentence and the second text
sentence on the basis of the calculated distance between the
first R tree and the second R tree, wherein:

the tree structure conversion section allots:

word information contained in the first text
20 sentence to the vertexes of the first R tree;

word information contained in the second text
sentence to the vertexes of the second R tree;

case information contained in the first text
sentence to the edges of the first R tree; and

25 case information contained in the second text

sentence to the edges of the second R tree.

14. The text sentence comparison apparatus according to claim 13, wherein the distance calculation section

5 calculates:

a mapping weight of a mapping from the first R tree to the second R tree,

a distance between a forest, which the first R tree includes, and a forest, which the second R tree include, and

10 a distance between a subtree, which the first R tree includes, and a subtree, which the second R tree includes.

15. The text sentence comparison method according to claim 14, wherein:

15 the distance calculation section calculates the mapping weight on the basis of a word substitution weight, a word deletion weight, a word insertion weight, a case substitution weight, a case deletion weight, and a case insertion weight.

20 16. The text sentence comparison method according to claim 13, further comprising:

a setting input section for allowing a user to set the condition of the mapping between the first and second R trees.

25 17. The text sentence comparison method according to

claim 13, wherein the condition of the mapping between the first and second R trees includes:

the mapping is a one-to-one mapping;
the mapping preserves parent-child relationship;
5 the mapping preserves structure; and
the mapping between the vertexes does not intersect with
the mapping between the edges.

18. The text sentence comparison method according to
10 claim 13, further comprising:

an output section for outputting the calculated distance
between the first text sentence and the second text sentence.

19. A text sentence comparison apparatus comprising:
15 an input section for inputting a first text sentence and
a second text sentence;

a tree structure conversion section for converting the
first text sentence and the second text sentence into a first
RO tree and a second RO tree, respectively;

20 a distance calculation section for calculating a distance
between the first RO tree and the second RO tree on the basis
of a distance between two RO trees, which is defined at least
in accordance with a condition of a mapping between vertexes
and edges of the two RO trees; and

25 a semantic content comparison section for calculating

a distance between the first text sentence and the second text sentence on the basis of the calculated distance between the first RO tree and the second RO tree, wherein:

the tree structure conversion section allots:

5 word information contained in the first text sentence to the vertexes of the first RO tree;

 word information contained in the second text sentence to the vertexes of the second RO tree;

10 case information contained in the first text sentence to the edges of the first RO tree; and

 case information contained in the second text sentence to the edges of the second RO tree.

20. The text sentence comparison apparatus according to claim 19, wherein the distance calculation section calculates:

 a mapping weight of a mapping from the first RO tree to the second RO tree,

20 a distance between a forest, which the first RO tree includes, and a forest, which the second RO tree include, and a distance between a subtree, which the first RO tree includes, and a subtree, which the second RO tree includes.

21. The text sentence comparison method according to claim 19, wherein:

the distance calculation section calculates the mapping weight on the basis of a word substitution weight, a word deletion weight, a word insertion weight, a case substitution weight, a case deletion weight, and a case insertion weight.

5

22. The text sentence comparison method according to claim 19, further comprising:

a setting input section for allowing a user to set the condition of the mapping between the first and second R0 trees.

10

23. The text sentence comparison method according to claim 19, wherein the condition of the mapping between the first and second R trees includes:

the mapping is a one-to-one mapping;

15

the mapping preserves parent-child relationship;

the mapping preserves brother relationship;

the mapping preserves structure; and

the mapping between the vertexes does not intersect with the mapping between the edges.

20

24. The text sentence comparison method according to claim 19, further comprising:

an output section for outputting the calculated distance between the first text sentence and the second text sentence.

25